Tétreault, É., Bélanger, M.-È., Bernier, A., & Carrier, J. (2018). Actigraphy data in pediatric research: The role of sleep diaries. *Sleep Medicine*, *47*, 86-92.

Abstract

Background: When assessing children's sleep using actigraphy, researchers usually rely on a sleep diary completed by a parent as an aid in scoring actigraphic data. However, parental non-adherence in completing the sleep diary may significantly reduce the amount of available data. The current study examined the agreement between actigraphic data scored with and without a sleep diary in order to evaluate the impact of not using a sleep diary when studying children's sleep with actigraphy.

Methods: Sixty children (6-10 years; 36 girls) wore an actigraph for three to seven consecutive nights, and mothers were asked to complete a diary of their child's sleep during the same period. Actigraphic data were scored using two conditions (with and without diary) rated independently for each child by two different research assistants, who each scored 50% of the files in each condition.

Results: Group-level analyses and intraclass correlations revealed very strong convergence between the two scoring conditions: on all sleep variables (sleep duration, wake duration, and sleep efficiency), average mean differences were very small and intraclass correlations very high. Bland and Altman's (1999) approach allowed for a child-by-child examination of agreement between the two conditions and revealed that, although they cannot be considered interchangeable, the two conditions produce quite minimal differences in the estimation of sleep variables.

Conclusions: The findings suggest that it is possible to use some actigraphy data for which no corresponding diary data are available, although this approach should be used sparingly.

Keywords: Sleep, Actigraphy, Children, Sleep diary, Scoring

1. Introduction

Children's sleep problems are common (Mindell & Owens, 2010) and usually persist throughout childhood (Lavigne et al., 1999). It has been observed that sleep difficulties tend to become chronic around 3 years of age, such that two thirds of children who had sleep problems before this age still suffered from such problems five years later (Thome & Skuladottir, 2005). Furthermore, sleep is linked to multiple areas of child development, such as cognitive (Bernier, Beauchamp, Bouvette-Turcot, Carlson, & Carrier, 2013), emotional (Gregory et al., 2005), behavioral (Bélanger, Bernier, Simard, Desrosiers, & Carrier, 2016) and physiological (e.g., later risk of obesity; Reilly et al., 2005) adjustment, suggesting that sleep disturbances can have a negative impact on several spheres of child functioning. Thus, accurate assessment of the duration and quality of children's sleep is important for pediatric research and practice.

Different methods are used to evaluate sleep in childhood; some are subjective (child or parent interviews, parental retrospective questionnaires, prospective sleep diaries), whereas others are objective. Among the latter, actigraphy, based on the use of a small watch-like computerized monitor, allows for non-intrusive assessment of sleep/wake patterns in infants', children's, and adolescents' natural environment (Sadeh, Lavie, Scher, Tirosh, & Epstein, 1991). The monitor can be worn at the wrist, the ankle or the waist. The number of movements per epochs (e.g., 30 sec) is recorded continually with a movement accelerometer. Scoring algorithms are then used to identify sleep and wake states from activity counts and to determine sleep variables such as sleep onset latency (number of minutes elapsed from bedtime to sleep onset), sleep duration (number of minutes scored as sleep between sleep onset and offset), frequency of awakenings (number of times that a predetermined number of minutes of wake occurs, e.g., 5 minutes), wake duration (number of minutes scored as wake between sleep onset and offset), and sleep efficiency (ratio of sleep duration to total time in bed *100). There are different brands of actigraphs and different scoring programs (e.g., Meltzer, Walsh, et al., 2012).

Studies suggest that some sleep variables are challenging to assess with actigraphy, as indicated by their poor convergence with corresponding polysomnographic (PSG) estimates (e.g., number of awakenings, sleep fragmentation or efficiency; Bélanger, Bernier, Paquet, Simard, & Carrier, 2013; Meltzer, Walsh, et al., 2012; Spruyt, Gozal, Dayyat, Roman, & Molfese, 2011). Overall, though, actigraphy is recognized as a valid and reliable method to assess sleep objectively (Berger et al., 2008; Meltzer, Montgomery-Downs, et al., 2012; Sadeh, 2011), as indicated notably by its convergence with PSG estimates on several sleep variables (e.g., sleep latency, sleep duration; Bélanger et al., 2013; Meltzer, Walsh, et al., 2012; Spruyt et al., 2011). Consequently, it has become frequent practice to use activity monitoring to assess children's sleep. Used jointly with validated algorithms, actigraphy constitutes a useful, valid, and relatively inexpensive way to collect naturalistic sleep data over multiple days.

There are, however, important challenges associated with actigraphic analyses, notably the detection of spurious motor activity counts associated with external motion (when a child sleeps in a vibrating vehicle like a car or a stroller) and the failure to detect motion due to removal of the actigraph. Furthermore, actigraphic algorithms cannot discriminate between motionless wakefulness (e.g., when the child is watching television in bed) and sleep (Sadeh, 2015). To improve detection of these situations, pediatric actigraphic studies usually rely on sleep diaries completed by a parent (often the mother). There are different types of sleep diaries, which vary in terms of the exact information asked from the parent. However, all sleep diaries can aid in scoring actigraphy data by helping to identify artifacts in actigraphy data, and also provide relevant information such as the approximate time when the child went to sleep and woke up the next morning, which helps determine sleep onset and offset times from the actogram (graphical representation of activity level provided by the actigraphy software). Diaries are also useful in identifying particular events that may have caused an atypical night for the child (e.g., illness or visitors at home). Researchers can then use this information in analyzing the data or to decide which data should be excluded (e.g., periods where the data are likely biased by artifacts). An important challenge, however, is that diary data are often not available. Although some do not provide the exact proportions, several papers using actigraphy mention that participants or specific nights were discarded given the parent's failure to fill out the sleep diary (e.g., Bélanger, Bernier, Simard, Bordeleau, & Carrier, 2015; Sitnick, Goodlin-Jones, Anders, 2008; Werner et al., 2008). In fact, a study by Acebo and colleagues (1999) found that when recording nights for a week with children and adolescents, a loss of up to 28% of data can be expected, for reasons including illness and technical problems, but also failure to complete the diary. This is not unexpected, as participants' perfect compliance with any research procedure can never be taken for granted. In the case of children's sleep diaries, the parent may fail to complete the sleep diary, on one or multiple days, and it is not developmentally realistic to expect children, especially young children, to assume responsibility for completing a sleep diary. While understandable, such non-compliance is of concern because even in cases where the diary was partially completed (e.g., on 3 or 4 nights out of 7), a participant's data could have to be deleted altogether, as Acebo et al. (1999) recommended that at least five nights of actigraphy recordings were necessary to assess sleep reliably. Thus, not only can parental (or child) non-adherence in completing the sleep diary reduce the number of scorable nights, it can also reduce the number of participants, due to the need to exclude participants who have fewer than a pre-specified number of scorable nights. Thus, non-adherence in completing the sleep diary constitutes a great challenge, as it may significantly reduce the number of nights and participants in a given study, potentially resulting in dramatic loss of statistical power. This is a salient issue for pediatric sleep studies, given that small samples are the norm rather than the exception in such studies. Furthermore, using only participants or nights for which sleep data are available on both measures (diary and actigraphy) could lead not only to reduced statistical power, but also to a non-random bias inherent to certain participants' or nights' characteristics, leading to less representative results.

One important yet under-studied question, then, is whether it might be possible to use some actigraphy data for which no corresponding diary data are available. With an adult population, Boudebesse et al. (2013) compared a no-diary condition with a diary condition, however only by looking at group-level average differences, which limits the conclusions that can be drawn. In what appears to be the only relevant study with a pediatric population, Meltzer and Westin (2011) evaluated the impact of using or not using a sleep diary to derive sleep actigraphy variables with the Micro Motionlogger Sleep Watch (Ambulatory Monitoring Inc.) and the Sadeh scoring algorithm. They concluded that while a sleep diary is essential to identify artifacts in actigraphic data, it may be possible to include one or two nights for which diary data are not available. They also indicated that additional research was necessary to examine scoring procedures using other devices and algorithms. To our knowledge, such additional research has yet to be conducted. Doing so is an important endeavor considering the increasing number of studies that use actigraphy to assess children's sleep, and the amount of data that can be lost due to failure to complete the sleep diary (e.g., Acebo et al., 1999).

Accordingly, the current study addressed the same question as Meltzer and Westin (2011), while benefiting from a larger sample (60 children, vs. 38) and using a different actigraphy device and algorithm (see section 2.3.1), which may lead to different conclusions (e.g., Meltzer, Walsh, et al., 2012). The aim of the current analyses was to compare two conditions of actigraphy scoring: a first condition in which the scoring of actigraphic data was guided by a parent diary, and a second condition in which the actigraphic data were scored without any use of a diary.

We also aimed to enrich the information drawn from the data in two ways. First, to estimate the degree of convergence between sleep estimates derived with and without the use of a diary for scoring, we used intraclass correlations (ICC) instead of Pearson's correlations. Like Pearson's correlations, ICC indicate the degree of relation between individuals' rank order on two measures (in this case, the two scoring conditions), however, ICC have the unique advantage of also considering systematic differences between measures (Brouwer et al., 2004; Weir, 2005 – more detail in section 2.5). Furthermore, we complemented the analyses with the Bland and Altman method (1999). In contrast to other statistical analyses (e.g., correlations examining degree of convergence, analysis of variance estimating mean-level differences), this procedure does not focus on between-children differences but rather estimates the agreement between methods on a child-by-child basis. This yields unique information that is masked by analyses based on measures of central tendency, notably the number of cases for which the discrepancy between two measures exceeds a pre-determined threshold considered to be meaningful. For this reason, Werner et al. (2008) stated that the Bland and Altman method constitutes the only appropriate approach to examine the agreement between two measures. In addition, such an analytic method may be uniquely informative in a clinical sense, owing to its focus on specific children's data points. Accordingly, we used this method to conduct a child-by-child examination of the agreement between actigraphic data scored with and without a sleep diary.

2. Method

2.1 Participants

Sixty typically-developing children (36 girls) aged between 6 and 10 years old participated in this study. In order to obtain a diverse and balanced age range, children were sampled into five equal groups, each corresponding to a distinct school level: first grade (M = 7.05 years; n = 12; 7 girls), second grade (M = 7.74 years; n = 12; 7 girls), third grade (M = 8.66 years; n = 12; 7 girls), fourth grade (*M* = 9.79 years; n = 12; 7 girls), and fifth grade (*M* = 10.54 years; n = 12; 8 girls). Families were part of a larger study and were recruited from birth lists randomly generated and provided by the Ministry of Health and Social Services. Written informed consent was obtained from parents prior to the beginning of the study. The study protocol was approved by the university's institutional review board. Criteria for participation were full-term pregnancy and the absence of any known physical and mental disability in the child. Most parents were Caucasian (80% of mothers, 75% of fathers). Mothers were between 20 and 45 years old (*M* = 31) and had 14.65 years of education on average. Fathers were between 22 and 48 years old (*M* = 33) and had 14.89 years of education on average. Family income varied from less than $20,000 CDN to over $100,000 CDN, with an average in the $60,000 to $79,000 bracket.

2.2 Procedure

At the end of a home visit, a research assistant left an actigraph and a sleep diary with the family and provided instructions concerning these two measures. Children put on the actigraph when they went to bed and removed it when they got up the next morning. They wore the actigraph on their non-dominant wrist for three (first and second grades) or seven (third to fifth grades) consecutive nights, and mothers were asked to complete a diary of their child's sleep during the same period. The three-night period with the younger children was chosen with the aim of reducing child and family burden and thus optimizing compliance. Although Acebo et al. (1999) recommended the use of at least five nights of actigraphy recordings to obtain optimally reliable measures, one notes from closer inspection of their data that satisfactory levels of reliability (close or superior to .70) can obtained with three days of assessment, which also guided our choice. More nights of assessment were nonetheless later added (third grade on), as children matured and were more amenable to wearing the actigraph for a full week while being able to remember to take it off for bathing and swimming, which reduced parents' burden.

Following the period of actigraphy assessment, the assistant returned to the family's home to retrieve the actigraph and the diary, and provided financial compensation to the family (CND$20 when the actigraph was worn for three nights and CND$30 for seven nights). Actigraphic data were then downloaded to a computer immediately upon returning to the laboratory.

2.3 Measures

2.3.1 Actigraphy

Children wore an AW-2 Actiwatch (Mini-Mitter, Philips Respironics, Bend, OR). Data were collected using the proportional integration mode (PIM) and scored using Actiware software version 6.0.9 (Philips Respironics, Bend, OR). The automated manufacturer's low sensitivity threshold algorithm (80 activity counts per epoch) was used to analyze data. This sensitivity threshold is frequently used to assess school-aged children's sleep (Meltzer, Walsh, & Peightal, 2015; Meltzer, Walsh, et al., 2012; Spruyt et al., 2011) and is well-suited to their enhanced motor activity during sleep (de Koninck, Lorrain, & Gagnon, 1992). Moreover, the AW-2 device, when used with this algorithm among children (aged 3 to 18 years; M = 8.8 years), shows good sensitivity (to detect sleep; 89%) and poorer but acceptable specificity (to detect wake; 77%) when compared with PSG estimates (Meltzer, Walsh, et al., 2012). Although non-optimal, this level of specificity is comparable to or higher than that obtained with other devices (e.g., Motionlogger Sleep Watch — Ambulatory Monitoring Inc.) and algorithms (e.g., Sadeh, Cole-Kripke, and high or medium sensitivity threshold algorithms), again with children from 3 to 18 years (see Meltzer, Walsh, et al., 2012). Although several actigraphy devices have an event marker that can be pressed by the participant to indicate bedtime and waketime, event markers were not used in the current study, given that they also often yield much missing data (e.g., nearly 30% of participants failed to press the event marker at sleep onset or offset in the study by Meltzer & Westin, 2011). This allowed us to investigate the validity of actigraphy data when scored without any supplemental information, whether from a diary or event marker. Data were collected in 30-s epochs. All first and second graders had three nights of valid sleep data. For third, fourth, and fifth graders, valid sleep data were available for seven nights for 20 participants, six nights for 11 participants, five nights for 1 participant, and four nights for 4 participants. Missing data were attributable to children's refusal or failure to wear the actigraph. Sleep diary data were available in all cases.

Sleep values were averaged across all nights available for each participant. Actigraphy-derived sleep variables were (a) nocturnal sleep duration, (b) nocturnal wake duration, and (c) sleep efficiency. Sleep onset and offset were determined as explained below.

2.3.2 Sleep diary

On the days when children wore the actigraph, their mothers completed a daily sleep diary (one diary by assessment period; i.e., period of three or seven days). Different sleep variables can be obtained from the diary (e.g., sleep duration, night awakenings perceived by the parent), but in the current study, the diary was used exclusively as an aid in scoring actigraphy data. For the first and second grade groups (in which children wore the actigraph for three nights), the diary consisted of a 24-hour timeline on which mothers indicated, for each half hour, whether their child was awake or asleep, and where he or she slept (e.g., child bedroom, car, etc.). Mothers were also asked to note the moments where the child did not wear the actigraph, as well as any event that might have disturbed the child's sleep, such as illness, holidays, or visitors staying late at home. With this diary, sleep onset (used to score the actigraphy data) corresponded to the time at which the mother noted the beginning of the first half hour of nighttime sleep, and sleep offset corresponded to the end of the last half hour of sleep, again as indicated by the mother. For the third to fifth grade groups (in which children wore the actigraph for seven nights), the diary consisted of a questionnaire on which the mother indicated, for each night of assessment, the time at which the child went to bed and the time he or she got up the next morning. With this diary, sleep onset corresponded to the time at which the mother indicated that the child went to sleep, and sleep offset corresponded to the time at which the mother indicated that the child got up the next morning. Mothers were also asked to note the moments where the child did not wear the actigraph and any event that might have disturbed his or her sleep.

2.4 Actigraphy scoring

For each participant, the two conditions (with and without diary) were rated independently by two different trained research assistants who were experienced with actigraphy data scoring. Each assistant was randomly assigned 50% of the actograms to score in each condition, with and without diary. For each age group, the first assistant scored 50% of the cases with the diary, and the other 50% without the diary. The second assistant did the exact opposite. Hence, each assistant scored a child's sleep only once, either with or without the diary (as they were never in charge of the same case in the two conditions), and the two assistants never scored the same child's data within the same condition. In order to score the data without any other cue (except the diary in the diary condition), the light indicator was removed from the actogram.

2.4.1 Diary condition

To determine sleep onset in the diary condition, the research assistant started examination of the actogram at 30 minutes prior to sleep onset indicated on the diary. Then, the assistant moved forward in the actogram from this starting point and manually defined sleep onset as the start of the first ten consecutive minutes of sleep (i.e., the first twenty 30-s periods for which the activity count was at or near zero). To determine sleep offset, the starting point was set 30 minutes after sleep offset indicated on the diary, and the research assistant moved backward in time from that point to define sleep offset as the end of the last ten consecutive minutes of sleep identified in the actogram. The 10-minute scoring rule is frequently used in pediatric sleep research (e.g., Kieckhefer, Ward, Tsai, & Lentz, 2008; Spruyt et al., 2011; Van der Heijden, Smits, & Gunning, 2006).

2.4.2 Without diary condition

In this condition, the diary was not used to score actigraphic data. In order to standardize the scoring of the two conditions and to avoid subjective judgment, the 10-minute scoring rule was also applied in this condition. Sleep onset was set at the beginning of the first ten consecutive minutes of sleep visually identified in the actogram (i.e., the first twenty 30-s periods for which the activity count was at or near zero) and sleep offset, at the end of the last ten consecutive minutes of sleep.

2.5 Data analysis

All statistical analyses were conducted using SPSS version 24 (SPSS Inc., Chicago, IL). For each sleep variable, we first averaged the nights available for each participant and conducted analyses of mean-level differences. We then computed intraclass correlations (ICC) for each sleep variable to estimate the degree of convergence between the diary and no-diary conditions. While Pearson's correlation (*r*) only measures the degree of association between two variables, ICC are designed to assess consistency or conformity between two measures or conditions. In fact, with Pearson’s *r*, any two measures may be highly (or even perfectly) correlated, and still differ systematically, as long as both measures increase or decrease at the same proportional rate. In contrast, ICC assess not only the strength of association, but also whether the measures differ systematically (Brouwer et al., 2004; Weir, 2005). Suppose that each child’s rank order relative to other children on sleep duration is the same across the two scoring conditions, but that the no-diary condition systematically yields higher (or lower) estimates of sleep duration than the diary condition, even by just a few minutes. This would yield a perfect (1.00) Pearson’s *r*, but a lower ICC, because ICC takes into account the systematic difference between the two measures. In contrast, if the difference between the two measures is random (with sleep duration sometimes higher with the diary, sometimes lower), then the ICC will be higher. Because an ICC is a correlation, it can vary from -1 to +1 and be interpreted like a Pearson's *r*.

In addition, the statistical approach proposed by Bland and Altman (1999) was used to assess the rate of agreement between diary- and without-diary-derived sleep variables. This approach allows for visual representation of agreement and differences between two measures or conditions by plotting them against each other. Bland-Altman limits of agreement for each variable are computed using mean and standard deviation of the differences between the two conditions. The recommended criterion to consider that two measures are interchangeable is that 95% of the data points lie within two standard deviations of the mean difference (limits of agreement = ±1.96*standard deviation). The data are graphed in a scatter plot with the Y axis presenting the difference between the two conditions (A-B) and the X axis the average of these measures ((A+B)/2).

3. Results

Descriptive statistics for the two conditions are shown in Table 1. For both conditions, sleep values were comparable to those obtained in prior studies using actigraphy with school-aged children (El-Sheikh, Kelly, Buckhalt, & Hinnant, 2010; Sadeh, Raviv, & Gruber, 2000). Two-way ANOVAs with sex and age group (first, second, third, fourth and fifth grades) as factors were performed on each sleep variable (sleep duration, wake duration, and efficiency; each in both conditions ─ with and without diary) and showed that there were no significant sex or age differences (or interactions) on any sleep variable (all *p*s >.11). Moreover, t-tests showed that there were no significant differences according to the type of diary (that used for first and second grades or that used for third to fifth grades); all *p*s >.06.

Table 1 shows that differences between sleep estimates derived with and without the sleep diary were extremely small on average: compared to the diary condition, the without-diary condition underestimated sleep duration by 1.01 minute per night and wake duration by 0.76 minute, and overestimated sleep efficiency by 0.11%. Nonetheless, t-tests were run between

Table 1
Descriptive statistics of sleep variables

Sleep variables	Min	Max	Mean	SD
Diary				
Sleep duration (min)	466.50	597.29	533.32	27.38
Wake duration (min)	9.79	64.50	32.50	10.63
Sleep efficiency (%)	89.42	98.17	94.27	1.80
Without diary				
Sleep duration (min)	466.75	581.86	532.31	25.89
Wake duration (min)	9.79	64.50	31.74	10.22
Sleep efficiency (%)	89.41	98.17	94.38	1.76

diary and no-diary conditions for each sleep variable and revealed that two differences were statistically significant (for wake duration, $p = .03$ and sleep efficiency, $p < .001$). Closer inspection of the data revealed that the differences between conditions, albeit very small in magnitude, were statistically significant because relatively systematic in direction: in 55/60 cases, sleep efficiency was higher when assessed without than with a diary, and wake duration was lower when assessed without than with a diary in 46/60 cases. These differences were further probed with the Bland-Altman analyses presented below. Intraclass correlation coefficients between the two conditions were very high (Landis & Koch, 1977) for all three sleep variables: ICC = .96 (sleep duration), .97 (wake duration) and .98 (sleep efficiency), all *p*s < .001.

Bland-Altman plots of the difference against the mean for all sleep estimates are presented in Figures 1 to 3. The mean differences obtained were normally distributed. In all three figures, circles represent children from the first and second grades group, and squares represent those from the third to fifth grades group, thus differentiating results obtained with the two different diaries. Figure 1 illustrates a comparison of total nightly sleep duration derived by the diary and without-diary conditions. Here, the mean difference between the two conditions was 1 min per

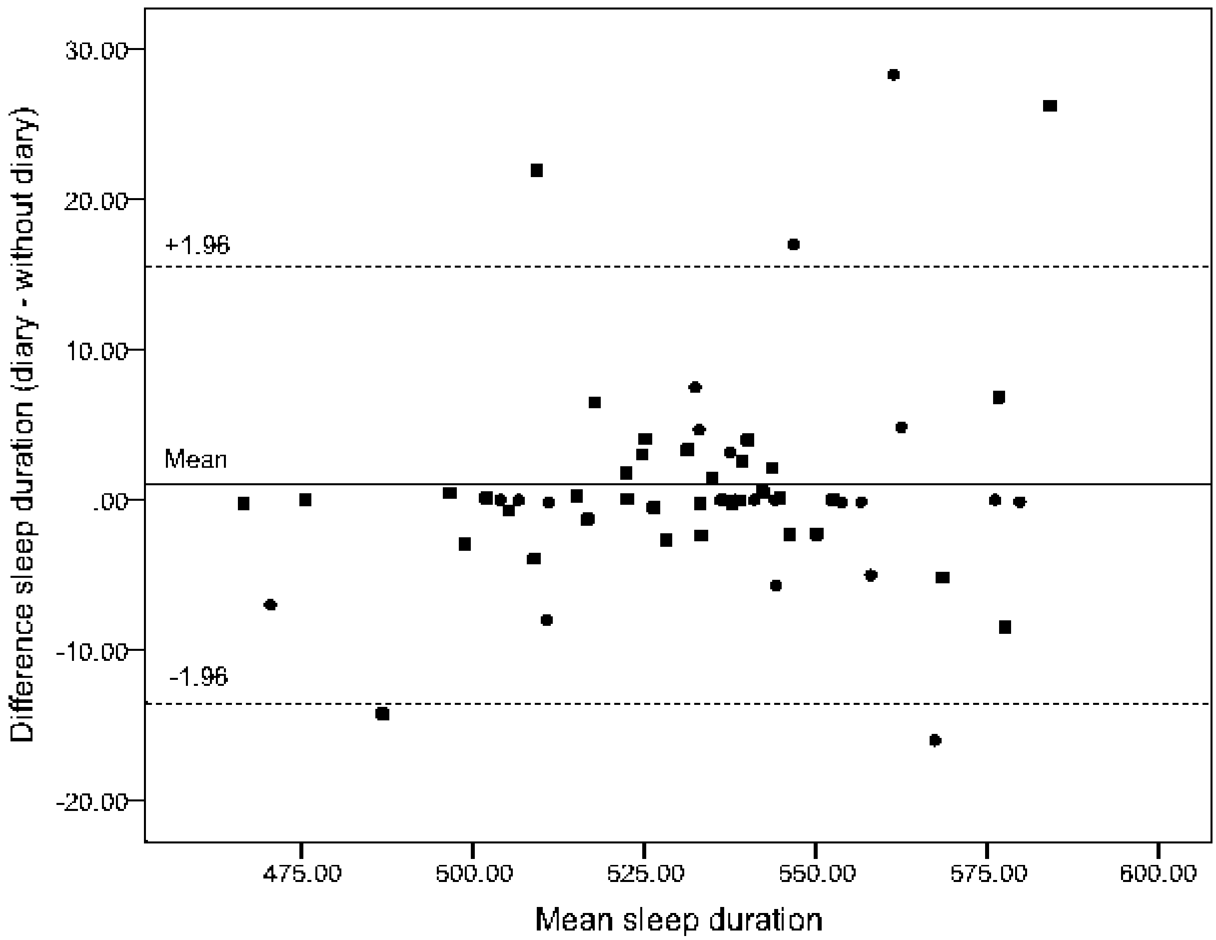

Figure 1. Plot of difference against mean for sleep duration. Circles represent children from the first and second grades groups (the first diary), and squares represent those from third to fifth grades groups (the second diary).

night [95% confidence interval: –55 s to +2 min 56 s]. Also, the difference in sleep duration for 90% of the children was between the limits of agreement, which were -13 min 32 s and 15 min

33 s. Similar results were found for wake duration as presented in Figure 2. The mean difference was 45 s [95% confidence interval 1 s to 1 min 25 s]. The limits of agreement were -4 min 16 s and 5 min 46 s and the difference in wake duration for 93% of the children was between these limits. Finally, Figure 3 shows a comparison of sleep efficiency measured by the diary and without-diary conditions. The mean difference was -.11% [95% confidence interval -.20% to -.02%]. The limits of agreement were between -0.80% and 0.59% and the difference in sleep

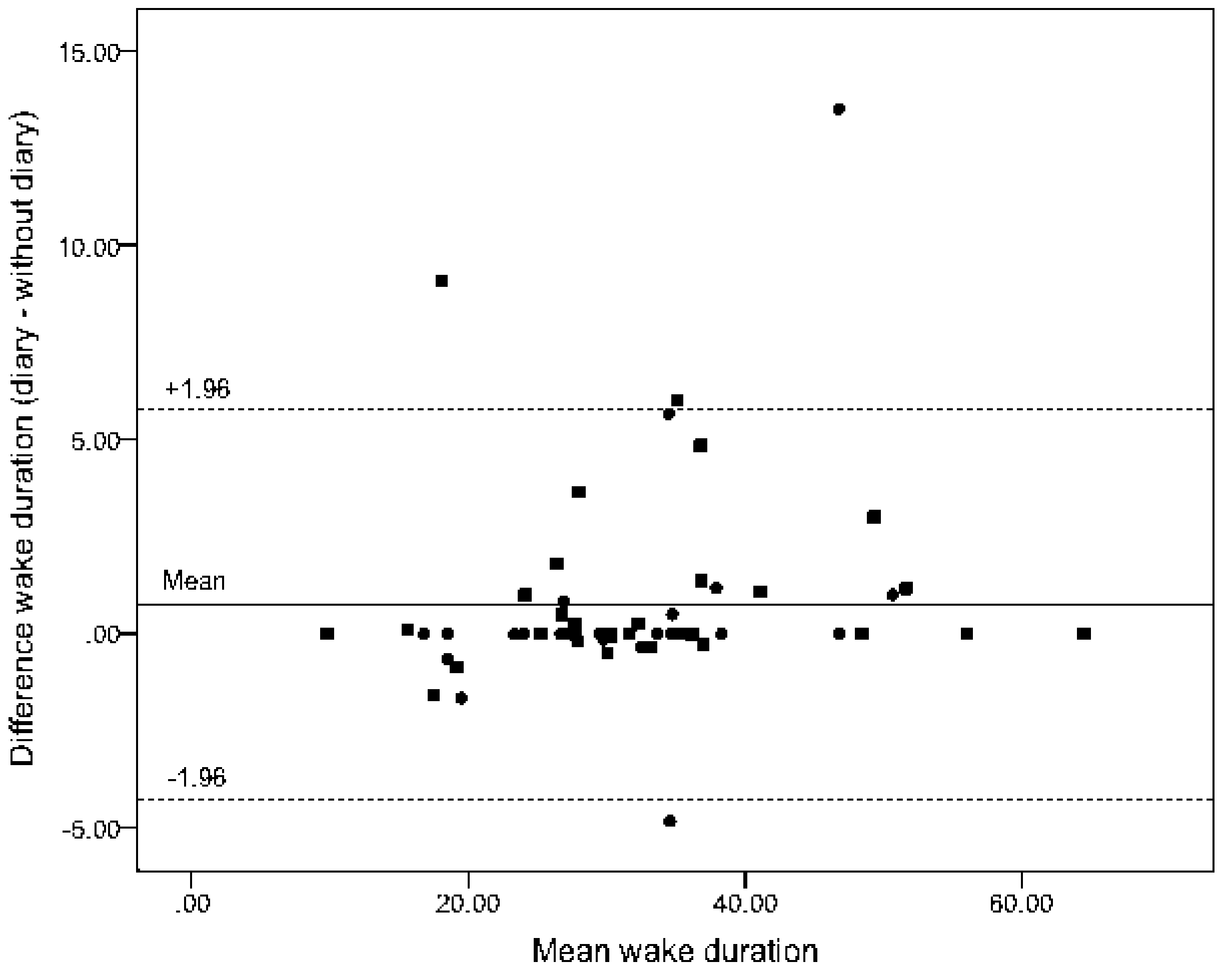

Figure 2. Plot of difference against mean for wake duration. Circles represent children from the first and second grades groups (the first diary), and squares represent those from third to fifth grades groups (the second diary).

efficiency for 93% of the children was between these limits. These 90% to 93% rates, while high, failed to meet the criterion that 95% of data must be within the agreement limits in order to confirm that two assessment conditions are *interchangeable* (Bland & Altman, 1999). However, for all children, the difference between the two conditions was less than 29 minutes for sleep duration, less than 14 minutes for wake duration, and less than 2% for sleep efficiency. Furthermore, inspection of the plots shows that the vast majority of difference scores were not

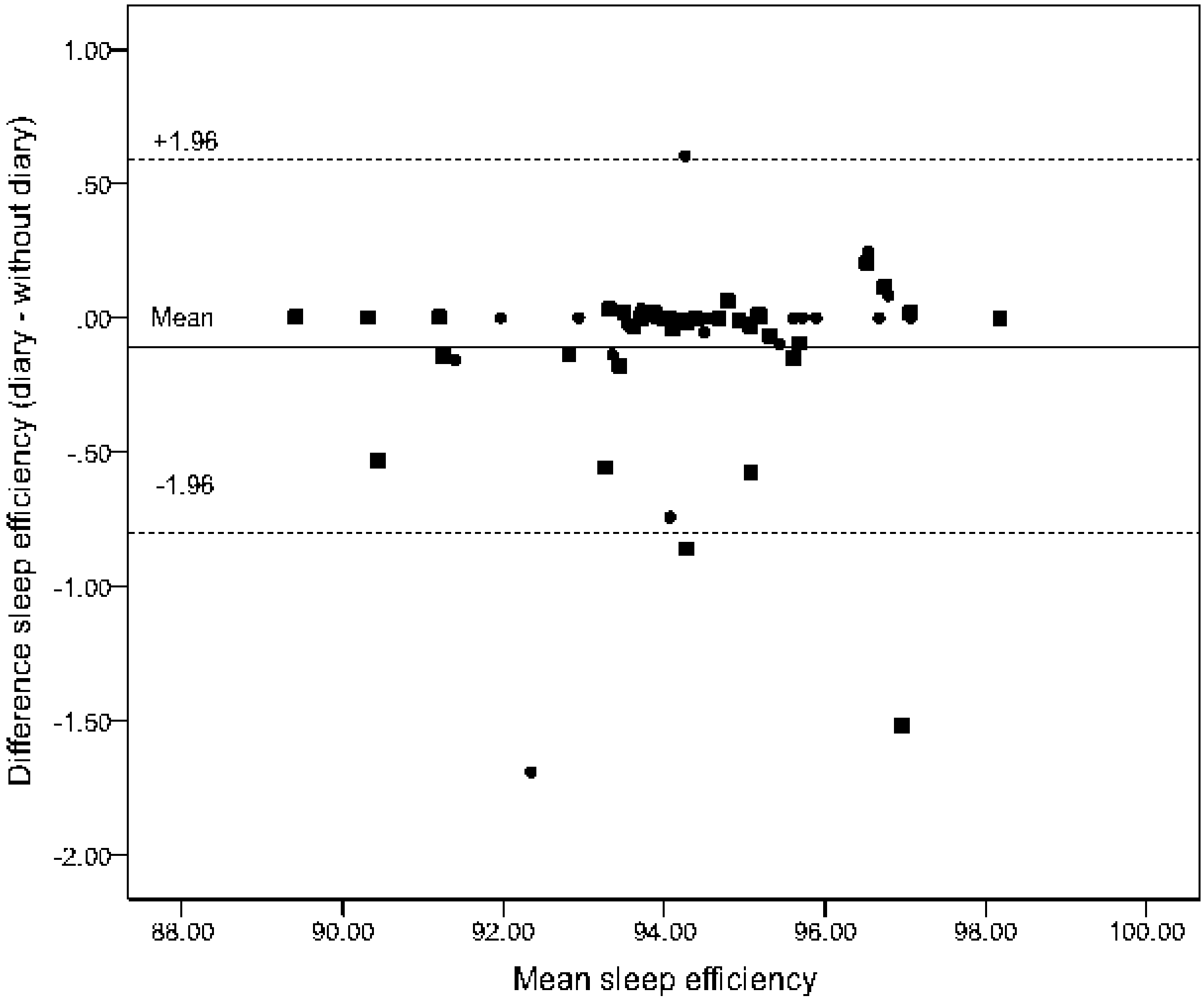

Figure 3. Plot of difference against mean for sleep efficiency. Circles represent children from the first and second grades groups (the first diary), and squares represent those from third to fifth grades groups (the second diary).

only within the limits of agreements, but in fact quite tightly distributed near the mean difference, which was very close to zero. The gaps between the two conditions were therefore quite small, even when analyzed on a child-by-child basis. The difference scores also did not appear to vary systematically with the mean scores, indicating that the discrepancy between the two measures was not related to mean levels of sleep duration, wake duration, or sleep efficiency. Finally, the plots indicate that outliers were not more likely to come from the use of one specific type of diary.

4. Discussion

The aim of the current study was to investigate whether children’s sleep and wake periods can be identified accurately with actigraphy when parents fail to complete a sleep diary. Sleep estimates derived from scoring the same actigraphic recordings with and without the support of a parental sleep diary were compared. To our knowledge, this is the first study to investigate this question with the AW-2 Actiwatch. Moreover, the chosen analyses allowed for the first investigation (to our knowledge) of the agreement between those scoring conditions not only at a group level, but also and importantly on a child-by-child basis.

The traditional group-level approach revealed very small differences between the two scoring conditions: the overall mean differences were 1.01 minute per night for sleep duration, 0.76 minute for wake duration, and 0.11% for sleep efficiency. Yet, these differences were statistically significant for both wake duration and sleep efficiency. As explained above, this was due to the relatively systematic *direction* of differences between scoring methods, which made for very small yet statistically reliable differences. The intraclass correlations were extremely high (Landis & Koch, 1977), varying from .96 to .98, indicating that children’s rank-order on all sleep parameters is virtually identical across the two scoring conditions.

The Bland-Altman analyses, which have not been used before to address such questions, revealed a more nuanced picture. Given the proportion (7-10%) of outliers whose values lied outside the limits of agreement, the data failed to demonstrate that the two conditions were *interchangeable* (Bland & Altman, 1999). However, for all children, the difference between the two conditions were quite small (less than 29 minutes for sleep duration, less than 14 minutes for wake duration, and less than 2% for sleep efficiency). Accordingly, the results suggest that the two conditions yield quite similar (albeit not interchangeable) results, even when analyzed on a child-by-child basis. In addition, the plots indicate that for a vast majority of children, the discrepancy between the two measures is much smaller, and this is so regardless of their sleep duration, efficiency, or wake duration.

Studies that assess children's sleep using actigraphy are generally mostly concerned with inter-individual differences, and the high intraclass correlations found in the current study show that, to this end, the two conditions (with and without diary) provide extremely similar results. In contrast, if one aimed to assess individual children's exact sleep or wake duration, for instance to establish precise norms or for clinical purposes, the small gap between the two conditions found in this study could justify the rejection of actigraphy data not corroborated by a sleep diary. Likewise, in sleep intervention studies using pre- and post-treatment sleep measures, the use of a sleep diary only at pre- or post-treatment could introduce some unwanted error likely to mask *or* inflate treatment effects. However, the current findings suggest that such error is likely to be small in magnitude in most cases.

Overall, the current analyses show a large degree of agreement between the two conditions, suggesting that it is possible to include some nights' or some participants' actigraphy data for which no corresponding diary data are available. The results are in line with those found by Meltzer and Westin (2011) using a different actigraph device and a different algorithm, also with school-aged children. However, we agree with Meltzer and Westin (2011) that analyzing child actigraphy data in the absence of a parental diary should be done with care. First, the diary contains relevant information for scoring, such as the approximate time when the child went to bed and woke up the next morning, which helps determine sleep onset and offset times from the actogram. Second, when scoring actigraphy data without a diary, potential errors remain, due for instance to artifacts that would not be detected in actigraphy data (e.g., a participant who removed the actigraph during the night or was sleeping in a car). On the other hand, discarding participants or specific nights on the basis of the parent's failure to fill out the sleep diary may lead to a non-negligible amount of lost data, especially due to the need to exclude participants who have fewer than a pre-specified number of scorable nights (e.g., five nights; Acebo et al., 1999). This may potentially result in decreased statistical power, as well as in a non-random bias inherent to certain participants' or nights' characteristics. We therefore recommend that researchers continue to ask parents to complete a sleep diary while their child is wearing the actigraph, bearing in mind that if necessary, modest amounts of actigraphic data that are not complemented by a diary can be scored validly and thus be included, especially in analyses of individual differences.

In the current study, the event marker was not used. However, it is worth noting that when available, the event marker constitutes a good replacement for the diary with respect to identification of sleep onset and offset. It must also be noted that when faced with actigraphy data not accompanied by any other information (diary, event marker, light indicator), some sleep variables may be impossible to assess (e.g., sleep onset latency, which was accordingly not considered in the current study).

The current study presents limitations that need to be considered when interpreting the results. First, the use of a low-risk and relatively small sample may have diminished variability in sleep patterns and led to low rates of children with sleep problems, making the results less generalizable to samples presenting with greater medical or psychosocial risk. In addition, some subjects had only three or four nights of sleep data, which contributes to reducing the amount of data analyzed and thus limits the chances of encountering particular situations that may jeopardize the validity of the sleep data, such as participants removing the actigraph.

5. Conclusions

The current report suggests that when studying typically-developing school-age children, it is possible to use a modest amount of actigraphy data not supplemented by a sleep diary, especially when conducting individual differences analyses. Yet, more studies, especially with clinical or high-risk populations, are necessary to further corroborate this suggestion and test its generalizability.

Acknowledgements: We gratefully acknowledge Natasha Whipple, Émilie Rochette, Natasha Ballen, Isabelle Demers, Jessica Laranjo, Célia Matte-Gagné, Marie-Ève Bélanger, Andrée-Anne Bouvette-Turcot, Marie Deschênes, Gabrielle Lalonde, Christine Gagné, Sarah Hertz, Marie-Soleil Sirois, Rachel Perrier, Élizabel Leblanc, Élodie Larose-Grégoire, Catherine Cimon-Paquet, and Nadine Marzougui for help with data collection.

Funding: This work was supported by the Social Sciences and Humanities Research Council of Canada [410-2010-1366]; the Fonds de Recherche du Québec - Société et Culture [2012-RP-144923]; and the Canadian Institutes of Health Research [MOP-119390].

References

Acebo, C., Sadeh, A., Seifer, R., Tzischinsky, O., Wolfson, A. R., Hafer, A., & Carskadon, M. A. (1999). Estimating sleep patterns with activity monitoring in children and adolescents: How many nights are necessary for reliable measures? *Sleep*, *22*, 95-103. Retrieved from http://europepmc.org/abstract/ med/9989370

Bélanger, M. È., Bernier, A., Paquet, J., Simard, V., & Carrier, J. (2013). Validating actigraphy as a measure of sleep for preschool children. *Journal of Clinical Sleep, 9*, 701-706. doi:10.5664/jcsm.2844

Bélanger, M. È., Bernier, A., Simard, V., Bordeleau, S., & Carrier, J. (2015). Attachment and sleep among toddlers: Disentangling attachment security and dependency. *Monographs of the Society for Research in Child Development, 80,* 125-140. doi:10.1111/mono.12148

Bélanger, M. È., Bernier, A., Simard, V., Desrosiers, K., & Carrier, J. (2016). Sleeping toward behavioral regulation: Relations between sleep and externalizing symptoms in toddlers and preschoolers. *Journal of Clinical Child & Adolescent Psychology*. Advance online publication. doi:10.1080/15374416.2015.1079782

Berger, A. M., Wielgus, K. K., Young-McCaughan, S., Fischer, P., Farr, L., & Lee, K. A. (2008). Methodological challenges when using actigraphy in research. *Journal of Pain and Symptom Management, 36*, 191-199. doi:10.1016/j.jpainsymman.2007.10.008

Bernier, A., Beauchamp, M. H., Bouvette-Turcot, A. A., Carlson, S. M., & Carrier, J. (2013). Sleep and cognition in preschool years: Specific links to executive functioning. *Child Development, 84*, 1542-1553. doi:10.1111/cdev.12063

Bland, J. M., & Altman, D. G. (1999). Measuring agreement in method comparisonstudies. *Statistical Methods in Medical Research*, *8*, 135-160. doi:10.1177 /096228029900800204

Boudebesse, C., Leboyer, M., Begley, A., Wood, A., Miewald, J., Hall, M., ... & Germain, A. (2013). Comparison of five actigraphy scoring methods with bipolar disorder. *Behavioral Sleep Medicine, 11,* 275-282. doi:10.1080/15402002.2012.685997

Brouwer, S., Kuijer, W., Dijkstra, P. U., Göeken, L. N., Groothoff, J. W., & Geertzen, J. H. (2004). Reliability and stability of the Roland Morris Disability Questionnaire: Intra class correlation and limits of agreement. *Disability and Rehabilitation, 26*, 162-165. doi:10.1080/09638280310001639713

de Koninck, J., Lorrain, D., & Gagnon, P. (1992). Sleep positions and position shifts in five age groups: An ontogenetic picture. *Sleep: Journal of Sleep Research & Sleep Medicine*, *15,* 143-149. doi:10.1093/sleep/15.2.143

El-Sheikh, M., Kelly, R. J., Buckhalt, J. A., & Hinnant, J. B. (2010). Children’s sleep and adjustment over time: The role of socioeconomic context. *Child Development, 81*, 870-883. doi:10.1111/j.1467-8624.2010.01439.x

Gregory, A. M., Caspi, A., Eley, T. C., Moffitt, T. E., O’Connor, T. G., & Poulton, R. (2005). Prospective longitudinal associations between persistent sleep problems in childhood and anxiety and depression disorders in adulthood. *Journal of Abnormal Child Psychology, 33*, 157-163. doi:10.1007/s10802-005-1824-0

Kieckhefer, G. M., Ward, T. M., Tsai, S. Y., & Lentz, M. J. (2008). Nighttime sleep and daytime nap patterns in school age children with and without asthma. *Journal of Developmental and Behavioral Pediatrics*, *29*, 338-344. doi:10.1097/DBP.0b013e318182a99e

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics,* 33, 159-174. doi:10.2307/2529310

Lavigne, J. V., Arend, R., Rosenbaum, D., Smith, A., Weissbluth, M., Binns, H. J., & Christoffel, K. K. (1999). Sleep and behavior problems among preschoolers. *Journal of Developmental and Behavioral Pediatrics, 20,* 164-169. doi:10.1097/ 000047

Meltzer, L. J., Montgomery-Downs, H. E., Insana, S. P., & Walsh, C. M. (2012). Use of actigraphy for assessment in pediatric sleep research. *Sleep Medicine Reviews*, *16*, 463-475. doi:10.1016/j.smrv.2011.10.002

Meltzer, L. J., Walsh, C. M., & Peightal, A. A. (2015). Comparison of actigraphy immobility rules with polysomnographic sleep onset latency in children and adolescents. *Sleep and Breathing, 19,* 1415-1423. doi:10.1007/s11325-015-1138-6

Meltzer, L. J., Walsh, C. M., Traylor, J., & Westin, A. M. (2012). Direct comparison of two new actigraphs and polysomnography in children and adolescents. *Sleep, 35*, 159-166. doi:10.5665/sleep.1608

Meltzer, L. J., & Westin, A. M. (2011). A comparison of actigraphy scoring rules used in pediatric research. *Sleep Medicine, 12,* 793-796. doi:10.1016/j.sleep.2011.03.011Mindell, J. A., & Owens, J. A. (2010). Evaluation of pediatric sleep disorders. In J.A. Mindell, & J.A. Owens (Eds.). *A clinical guide to pediatric sleep. Diagnosis and management of sleep problems*, (2nd ed., pp. 30-42). Philadelphia, PA: Lippincott.

Reilly, J. J., Armstrong, J. Dorosty, A. R., Emmett, P. M., Ness, A., Rogers, I., . . . for the Avon Longitudinal Study of Parents and Children Study Team (2005). Early life risk factors for obesity in childhood: Cohort study. *British Medical Journal, 330*, 1357–1363. doi:10.1136/bmj.38470.670903.E0

Sadeh, A. (2011). The role and validity of actigraphy in sleep medicine: An update. *Sleep Medicine Reviews, 15*, 259-267. doi:10.1016/j.smrv.2010.10.001

Sadeh, A. (2015). Sleep assessment methods. *Monographs of the Society for Research in Child Development*, *80*, 33-48. doi:10.1111/mono.12143

Sadeh, A., Lavie, P., Scher, A., Tirosh, E., & Epstein, R. (1991). Actigraphic home-monitoring sleep-disturbed and control infants and young children: A new method for pediatric assessment of sleep-wake patterns. *Pediatrics*, *87*, 494-499. Retrieved from http://pediatrics.aappublications.org/content/pediatrics/87/4/494.full.pdf

Sadeh, A., Raviv, A., & Gruber, R. (2000). Sleep patterns and sleep disruptions in school-age children. *Developmental Psychology*, *36*, 291-301. doi:10.1037/0012-1649.36.3.291

Sitnick, S. L., Goodlin-Jones, B. L., & Anders, T. F. (2008). The use of actigraphy to study sleep disorders in preschoolers: Some concerns about detection of nighttime awakenings. *Sleep*, *31*, 395-401. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2276752/

Spruyt, K., Gozal, D., Dayyat, E., Roman, A., & Molfese, D. L. (2011). Sleep assessments in healthy school-aged children using actigraphy: Concordance with polysomnography. *Journal of Sleep Research, 20*, 223-232. doi:10.1111/j.1365-2869.2010.00857.x

Thome, M., & Skuladottir, A. (2005). Changes in sleep problems, parents distress and impact of sleep problems from infancy to preschool age for referred and unreferred children. *Scandinavian Journal of Caring Sciences*, *19*, 86–94. doi:10.1111/j.1471-6712.2005.00322.x

Van der Heijden KB, Smits MG, Gunning WB (2006). Sleep hygiene and actigraphically evaluated sleep characteristics in children with ADHD and chronic sleep onset insomnia. *Journal of Sleep Research*, *15*, 55-62. doi:10.1111/j.1365-2869.2006.00491.x

Weir, J. P. (2005). Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *Journal of Strength and Conditioning Research*, *19*, 231-240. doi:10.1519/15184.1

Werner, H., Molinari, L., Guyer, C., & Jenni, O. G. (2008). Agreement rates between actigraphy, diary, and questionnaire for children's sleep patterns. *Archives of Pediatrics & Adolescent Medicine, 162*, 350-358. doi:10.1001/archpedi.162.4.350